DOCUMENT RESUME

ED 178 551 TM 008 945

AUTHOR Anderson, Scarvia B.
TITLE Competency Testing: Some Unlearned Lessons.
PUB DATE Mar 79
NOTE 18p.; Paper presented at the Michigan School Testing Conference (Ann Arbor, MI, March, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Academic Standards; Basic Skills; Definitions; Educational Accountability; Elementary Secondary Education; Examiners; Fundamental Concepts; *Individual Differences; Interpersonal Competence; *Minimum Competency Testing; Negative Attitudes; *Political Influences; *Testing Problems; Testing Programs

ABSTRACT

In a speech on minimum competency testing, the author discusses four issues which are still unresolved. The first is the need to recognize the political overtones which gave impetus to the competency movement as a reaction against the perceived failures of the schools and teacher education institutions, rather than as a concern for student development and academic growth. The second issue is the lack of a clear definition of educational or social basic competency, due to different operational definitions of such concepts as affect vs. cognition; proficiency vs. performance; and knowledge vs. application. Thirdly, there is a need to recognize the attempt at egalitarianism which ignores individual differences, talents, or handicaps. Alternatives to a common set of standards in competency testing programs may be helpful. The fourth issue is the defensive stance taken by those involved in testing toward setting standards and teaching to the test, in the context of teaching students the skills and knowledge measured by the competency tests. Instead, it is important to recognize reality and to help educators design tests which are worth teaching and learning for. (MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED178551

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY Scarvia Anderson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM 008 945

Competency Testing: Some Unlearned Lessons

Scarvia B. Anderson
Educational Testing Service

In 1971, I made several speeches on educational accountability which my mother called the "hot pants of education." Well, if accountability was the hot pants of education, then criterion-referenced tests were the sleek, knee-high boots--worn sometimes appropriately in Michigan's snows but more often inappropriately in Ft. Lauderdale or over-heated offices. And what does that make minimum competency testing? How about designer jeans? Purchasable only in the most exclusive department of J. L. Hudson, embroidered boldly with Calvin Klein or Gloria Vanderbilt across the rear, but still the same old denim. The minimum competency tests, in spite of some of their fancy names and claims, are by and large the same old objective tests of reading, English mechanics, and arithmetic--the Levis, if there are such, of educational testing.

I do not want to suggest that, just because an educational idea acquires a certain faddishness, there is something wrong with it. Indeed, accountability was a necessary concept for education to take on when it did. It gave voice to the growing concern about the degree of responsibility that public schools should have toward those who were footing the bill, and it continues to be the vehicle for many useful exchanges between educators and others who view the amount and nature of that responsibility differently.

The criterion-referenced camp can be credited, if for nothing else, with greatly improving the quality of achievement tests, by insisting on detailed specifications of the knowledge and skills to be measured. Thus, both instructors and students are provided with a clearer idea of what they are about when they give, take, or prepare for such tests. Professional test makers had long stressed the need for good specifications as the first step in test development. However, this advice frequently went unheeded at state and local levels, where tests were put together any old way or pulled

---

Speech presented at Michigan School Testing Conference, Ann Arbor, March, 1979.

off the shelf without regard for the match between the content of the test and the curriculum. The importance of the contribution of the criterion-referenced testers to improving achievement test quality cannot be underestimated, because achievement tests, not aptitude tests, are the measures that have a real role to play in the educational process.

Minimum competency tests are one kind of achievement test. Unfortunately they are usually a rather trivial kind with respect to content, although they may be far from trivial in some of their applications. As all of you who read Chris Pipho's (1978) reports know, almost all of the states have some kind of minimum competency testing requirement in place, proposed, or at least under study. There are also some local school boards who independently have set up their own requirements. Those on whom such requirements actually fall are most often high school juniors and seniors or others who are hoping to escape the school environment. If they pass the tests, they can get away early or honorably.

Your Conference planners--Dr. Womer, to be specific--have asked us to comment today about what we have learned from all of this competency testing activity. The main lesson I have learned--too late, for the programs were already printed--was that anyone was crazy to agree to give yet one more paper on this topic. I learned this simply by carrying around a little bit of the competency testing literature. It was very heavy. It also seemed to be very complete. However, I couldn't be sure, for there is no way that a busy person could have got through everything in my briefcase, let alone the great bulk waiting back on the bookshelves. Therefore, I am sure that I have missed some terrific insights and, thus, shall probably draw some quite erroneous conclusions. But please be tolerant for a few minutes now; I understand that there will be time for you to straighten me out later.

I am sure that a lot of people have learned a lot from the experiences of the last few years in competency testing. For example, some state testing directors have learned how to talk more easily with the press, legal counsel, and members of the state legislature. Test developers have learned not to say that you don't need to pretest items. School superintendents, especially in districts with powerful unions, have learned that they can't make personnel decisions about teachers on the basis of student performance on competency tests. Departments of education have learned that large-scale competency testing programs cost a lot of money, as do any remedial efforts geared to the results. Teachers have learned that students will complain if what's on the tests is not what they've been taught. And, certainly, professors in education and the social sciences who are under pressure to publish have learned that articles on "problems with competency testing" are likely to find quick favor with editors. All of these are important lessons, and I do not want to minimize them.

But I think there are some other very important lessons that we have not learned, at least to any degree of mastery, and I would like to discuss four of them today. They do not quite fall into the same category as the list of negatives, ills, or problem areas associated with minimum competency testing to which everyone from Gene Glass (1978) to the American Friends Service Committee (1978) has contributed. Rather they are more in the nature of understandings we need to come to in our pursuit of the best way to assess competency.

1. The first lesson is on "the games people play." The minimum competency testing movement is rooted in negative, distrustful, sometimes punitive, and very political soil. This is true, of course, for many of

the major social and educational movements in our history; furthermore, change has proceeded most rapidly, if not rationally, when the impetus was angriest. The competency testing movement took hold very quickly indeed, when we consider how long it usually takes for anything significant to happen in education. But the SAT scores had gone down. All but the most selective colleges found it necessary to put in more remedial courses --and some of the selective ones reinstituted old basic requirements like freshman English. Parents who had prided themselves on not interfering with the education of their children--leave that to the professionals-- suddenly decided to turn off the TV for a minute and listen to how well Johnny could read, and then felt put upon if it wasn't very well. Minority groups, lulled briefly in the security of equality of educational opportunity, started questioning what that opportunity was for--or worth. Employers, operating under considerable pressure from government agencies and the courts to show a greater diversity of employees on their rolls, decried the lack of basic skills in their new hires. Of course word got to the legislators. The schools were the villains--the teachers (especially if they had fought and won tenure, fair wages, or reasonable working conditions), the "new" curricula and "watered-down" textbooks, the administrators (who were almost too busy at first with court actions, failing bond issues, violence, and vandalism to notice that they were under attack), and those who trained the teachers and the administrators in state colleges and universities that used to be called teachers colleges.

Minimum competency testing, as practiced today, has more than a little "get the schools" motivation behind it and a lot less concern for student development than many of its proponents will admit. In spite of this, competency testing was probably a good idea and the programs can become a positive force for educational improvement. However, it has been a

mistake to forget or pretend to forget its roots, for then we argue its merits on false premises and try to answer what are essentially political questions with puffy professional pronouncements.

2. The second lesson that we haven't learned well enough is the "I know you believe you understand what you think I said, but what you heard is not what I meant" lesson. I am willing to bet that no two persons in this room would define educational or social competency in exactly the same way. We know that even when people come to some reluctant agreement about the nature of a competency domain, they are likely to part company again over the issue of what represents minimum competency in that domain. In Florida, competency means functional literacy in language and mathematics; in Virginia, they add citizenship, but it is probably a rather different citizenship from that in Pennsylvania. In New Jersey, The Bilingual Minimal Standards Committee is defining minimum basic skills in terms of "native language communication skills, second language communication skills, and native and second language computational skills" (Burke, 1978). And, in Texas, a recent newspaper story said there was continued wide support for high school football and marching bands as skills that are basic in that state. Jarrett summed it up nicely in the title to his 1977 paper in the Phi Delta Kappan: "I'm for Basics, But Let Me Define Them."

Messick and I (1974), with a lot of egging on by the usual panel of experts and what was then the Office of Child Development, had the temerity to describe twenty-nine facets of social competency in young children. These ranged from "realistic appraisal of self, accompanied by feelings of personal worth," to "appropriate regulation of antisocial behavior," to "curiosity," "fine motor dexterity," "quantitative and relational concepts, understandings, and skills," "enjoyment of humor, play, and fantasy," and

"competence motivation" (you not only want the child to be competent, you want him to want to be competent). Think what a list we might have come up with if we had given similar free rein to describing the facets of competency for 16-year-olds.

Bracey (1978), with appropriate reference to psychiatric theory, warns of the dangers of contemporary separation of affect and cognition. He also reminds us of the conceptual and practical problems we run into when we confuse competence and competent behavior, or, as Messick and I labeled them, proficiency and performance. Affect vs. cognition, proficiency vs. performance--those are only two of the issues we get hung up on, trying to define and measure competence. Some have as the focus of their competency assessment, skills deemed necessary for survival in society; others stress skills needed for survival in school and progressing to the next academic level. Of course, survival in school and survival in society are not necessarily the same thing, although, as Haney and Madaus (1978) point out, we probably don't know what the important life skills are.

Still other differences that characterize people's operational definitions of competency revolve around knowledge vs. application, maximal vs. typical performance, and "positive components of competency (the characteristics we can agree we want the person to have or develop) ... [vs.] negative characteristics which may serve as obstacles to learning, development, or societal adjustment" (Anderson & Messick, 1974).

Finally, some definitions of competency are intrinsically tied to practical limitations in test design and use. Achievement tests are basically of two types: those where the student recognizes the appropriate response and those where the student constructs the appropriate response.

It is not uncommon to hear school administrators say that they are measuring writing skills when in fact they are measuring students' ability to recognize correct forms, or that they are measuring laboratory skills when the tests require students simply to recognize appropriate laboratory procedures from word and picture descriptions. Even if students make high scores, we cannot be confident in the first example that the student, himself, can write acceptably or, in the second example, that the student might not add water to sulphuric acid instead of vice versa, when left on her own in the lab.

In spite of all of these variations in what is meant or implied when people talk about competency, we tend to praise it, rail against it, philosophize about it, and attempt to improve it, as if "it" were a unitary thing. "It" ain't. And, as long as we go on treating competency testing that way in our journals, professional meetings, and especially in conversations with lay persons, confusion will reign. Ultimately, as we used to say in the old days about intelligence, competency is what the tests measure--plus the other things we couldn't figure out how, or afford, to measure. If we get our definitions clear in the first place, we'll have a better chance to measure what we can measure better and at least have a good and communicable picture of what we've left out.

3. Third lesson: the "every little grain of sand" lesson. Most of us were trained in disciplines, or disciplines derived from disciplines, that expected, described, and sometimes cherished human differences. Yet the minimum competency movement, if taken at face value--as it is by so many legislators and journalists--has in it the suggestion of an egalitarianism that even the radical sociologists might object to (but see Britell, 1978).

We learned some years ago that words like "meritocracy" and "mental deficiency" turned many readers and listeners off. But this did not stop us from continuing to notice and nourish those who seemed to have significant talents and potential in art, music, science, mathematics, mechanics, and other fields of endeavor. Advanced placement programs have been thriving in recent years; Julian Stanley's (1977) work with mathematically precocious youth has received international attention; the Reverend Jesse Jackson doesn't want his followers to be content with mediocre skills--no way. Many individual public school districts--Flint is an example--devote considerable energy to identifying and providing special programs for those with the talent and motivation to achieve at a level above average (and even "average" is well above "minimum" in most connotations). It is a rare school indeed that does not point with pride to its National Merit Scholar, athletic star, poetry contest winner, and graduate who made it on the rock circuit, as a college professor, or in real estate. Of course, the most frequently given minimum competency tests don't pretend to get into many of these areas, and their defenders protest that the measures are intended to cover only a few fundamental and universally needed skills in a limited way. "Being able to perform reasonably well on the test tasks is necessary," they say, "but nobody ever said it was sufficient--there's a lot more to education than that." Yet simply having minimum competency tests without equally well publicized accounts of efforts to encourage and measure that "lots more" leaves the impression that something below mediocrity is the goal.

The same period that has seen the rise of the competency testing movement has also been marked by extraordinary public, legislative, and

legal concern for the handicapped. The rule of the day is to "mainstream" those with mental, physical, and emotional handicaps in as many subjects and social and athletic activities as possible. The success of such a drastic shift in attitudes about the best way to educate the handicapped has yet to be demonstrated and will rest, we suspect, as much on the system for identifying those to be mainstreamed as on the treatment they get after they are placed in "regular" classes and activities. However, their presence in the schools gives special salience to the question of whether a school system should have a common set of minimal competency demands and standards for everyone. I have seen very little in the descriptions of minimum competency programs about what procedures are followed with students who are blind, palsied, mentally handicapped, or otherwise unlikely to be able to respond to the test materials as intended. One state suggests that the reading test passages and questions be read aloud to blind students, changing the skill measured rather substantially, I would think. The Maryland legislation pertaining to minimum competency testing specifies that "all students except the moderate and the severe and profound intellectually limited...be included" in the specified grades, although the objective bases for exclusion were not spelled out in the account I read (Pipho, 1977). If students are excluded from the testing requirements for whatever reason or the testing procedures are changed for them, what explanation then goes on their diplomas? What kinds of statements <u>can</u> be made--legally?

To establish a common set of competency requirements for all those who want to be telephone operators in Detroit, play on the professional golf circuit, enter the Rachmaninoff piano competition, or receive a

Ph.D. in clinical psychology from Michigan State seems reasonable. Each of the groups, through self-selection, is already relatively homogeneous in a number of respects--certainly in comparison with the population in general--and all of the members have similar aspirations. The public schools, on the other hand, capture almost the full variation of human characteristics present in the total population, and the students leaving any one of those schools are heading off in almost as many directions as their numbers. These conditions give a substantially different meaning to the notion of a common set of requirements from that implied in the examples about telephone operators, golfers, pianists, and Doctoral candidates. No, I am not suggesting that we must throw out the concept of minimum competency in education. But the ideological schizophrenia I mentioned at the beginning of this discussion of lesson 3, as well as the feeling of discomfort that many have with <u>any</u> educational process that requires adherence to a common set of standards in the face of known tremendous variability in the characteristics and prospects of the group to whom those standards are applied, strongly suggest a re-examination of any commitments to the same competency requirements for everyone. Thoughtful attention to alternatives may be salutary both for the competency testing movement and for public education in general.

4. The last lesson is "the best defense is a good offense." Or why are we letting those people who don't know much about tests and measurements beat up so on the educators and social scientists who <u>do</u>? I think that it's at least partly because we don't take enough time to try to find out what their concerns really are and enough initiative in telling them forthrightly what tests can and can't do for them, in language they can understand without a dozen equivocations in every paragraph. We wait until they have

us firmly in a corner and then try to answer their reasonable enough questions or counteract their seemingly logical charges with a lot of professional double talk. Let me illustrate what I mean by reference to two areas: standard setting and teaching to the test.

We're mighty defensive about standard setting. We do a lot of clever infighting about whether Nedelsky or Ebel or Angoff is right. Probably all of them are "right," depending on the nature of the judgments about test content that you want to make. The methods answer somewhat different questions, and each has its strengths and weaknesses. Furthermore, in our infatuation with criterion-referenced tests and content validity, we are overlooking situations where normative information and empirical validity studies may be more informative to reasonable decisions about standards and cut-off scores. If we have 25 places in the advanced mathematics course and 100 applicants, a content validity approach isn't going to do us much good. With such candidates, we might very well find that 80 or more perform sufficiently well on an appropriate achievement measure that we have no reason to doubt their capability to function quite satisfactorily in the advanced course. At this point, we might want to take our own advice and use other indicators in addition to the scores on the achievement test--probably grades in previous math courses, maybe teacher nominations. But, combining these data, we may still have an eligible list of 50 or so. And, unfortunately, as rational and fair as the suggestions to choose randomly from the qualified group seem (e.g., see Lerner, 1978), it is impossible to get parents, students, or even teachers to buy such a procedure. So we might as well take the top 25 candidates in terms of the combined ratings and be done with it.

The main point is that even though we know that the procedures we suggest for setting standards are less than perfect, they are usually far

better than what schools have been doing. And schools have been in the business of setting and enforcing the most arbitrary standards for years. Remember the teacher who flunked a student in his course because the student's average exam grade was 69 1/2? And the student who failed to graduate because he lacked one term of phys. ed.?

"Teaching to the test"--most measurement types say it's a bad thing. Most teachers do it. If teachers don't do as much of it as parents think they should, then parents who can afford it put their children in other schools or send them to special coaching schools. No tests and measurements expert in the world stands a chance of getting a parent to back down when that parent says, "I want Maria or Gary to study for that test. You can't tell me that studying for a test isn't a good thing to do." It's true that some kinds of tests, especially the sophisticated verbal aptitude tests, are harder to prepare for, and especially to cram for, than other kinds of tests. But we are not talking about aptitude tests here; by and large when we talk about competency testing we are--or should be--talking about achievement tests. And if students acquire the skills and knowledge measured by the tests, then that's all to the good. The problem arises if students are taught the answers to the very test they will be given later to appraise their competency or determine their promotability. Especially if some teachers and students have access to the test and the answers and others do not. However, we can think of some situations where there is nothing wrong with teaching students exactly what they will be asked on a test--for example, the multiplication table, the symbols for the elements in the periodic table, how to transpose a given melody from the key of C to C minor, or short forms in shorthand.

I'll tell you a story about the director of research in a large Southern city. He was dismayed because the students' scores on a national standardized

test had dropped from the year before, and he wouldn't release them. First, he got the computer center to rescore the tests--he was sure there was an error. There wasn't. Then he had them check on the equating to the earlier form. "OK," they said. Then he had them re-equate going the other way. Still, OK. Finally, after several weeks of this and clamors on all sides to see the scores, the head of the computing center met the research director in the parking lot. "Hey, buddy," he said, "I think I've figured out how to get those test scores up." "How? How?" insisted the research director, with relief showing on his face for the first time in weeks. "Teach those little bastards how to read," was the answer.

Teachers are going to teach to tests. Students are going to learn for tests. Let's recognize these facts and design, or help teachers, schools, and state departments design, tests that are worth teaching and learning for.

. . .

To steal a story from Dean John Cribbet (1979):

> A leading educator recently gave a speech on the subject "Our Two Greatest Perils: Ignorance and Apathy.: One listener turned to another and asked, "What did he say?" Came the response: "I don't know and I don't give a damn."

For those of you who would like to know what I said, it was that, although we have made progress in dealing with the technical and practical problems associated with large-scale competency testing efforts, there are still some major lessons we need to apply ourselves to.

1. The first, which I dubbed the "games people play" lesson, has to do with recognizing the essentially political nature of the competency testing movement in the schools. Our aim as tests and measurements persons should be to contribute to the process of making competency testing a real force for educational improvement. But rationality cannot be imposed; it will have to evolve.

2. The second lesson is derived from that saying you see so often on gift shop items: "I know you believe you understand...." We know that one state's or school's or person's definition of "competency" is not another's, but we tend to slip into treating all competency testing programs as if they were the same thing, with the same problems and issues. Before we criticize or try to help a program, we must strive for a clear understanding of the domains to be measured and <u>why</u>--whether they are termed "basics" or represent a more ambitious array.

3. Third--the "little grains of sand" lesson--I urged a reexamination of the widely held proposition that competency demands should be the same for all the members of a given student population, especially in the face of known exceptions. At the least, there should be as explicit a rationale for common standards as for the measurement domains.

4. Finally, I asked if we've learned the lesson that "the best defense is a good offense" with respect to the criticisms of testing and test-related procedures that we all face. True, we need to spend more time trying to understand exactly what the users' problems are and how we can best help them. And we certainly need to be wary of any exaggerated claims about what tests can do. But, beyond that, let's be proud of the help we can provide education through measurement and communicate that pride and our expertise in certain terms.

Now for those of you who "don't give a damn," I hope that applies to my speech and not to the challenges of competency testing.

References

American Friends Service Committee. Southeastern Public Education Program: A citizens' introduction to minimum competence programs for students. Columbia, S.C.: American Friends Service Committee, 1978.

Anderson, S., & Messick, S. Social competency in young children. Developmental Psychology, 1974, 10, 282-293.

Bracey, G. W. Some reservations about minimum competency testing. Phi Delta Kappan, 1978, 59, 549-552.

Britell, J. K. Competence and excellence: The search for an egalitarian standard; the demand for a universal guarantee. Address presented at the Topical Conference on Minimum Competency Achievement Testing of the American Educational Research Association in Washington, D.C., October 1978.

Burke, F. G. Memo from Commissioner Burke to members of the New Jersey State Board of Education, December 6, 1978.

Cribbet, J. E. Presidential Address, Association of American Law Schools; quoted in Marginalia, The Chronicle of Higher Education, January 22, 1979, 17, (19), 2.

Glass, G. V. Minimum competency and incompetence in Florida. Phi Delta Kappan, 1978, 59, 602-605.

Haney, W. E., & Madaus, G. Making sense of the competency testing movement. Harvard Educational Review, 1978, 48 (4), 462-485.

Jarrett, J. L. I'm for basics, but let me define them. Phi Delta Kappan, 1978, 59, 235-239.

Lerner, B. Equal protection and external screening: Davis, De Funis, and Bakke. Proceedings of the 1977 ETS Invitational Conference. Princeton, N. J.: Educational Testing Service, 1978.

Pipho, C. Update VII: Minimal competency testing (Report No. 105). Denver, Colo.: Education Commission of the States, 1977.

Pipho, C. State activity minimal competency testing. Denver, Colo.: Education Commission of the States, May 1978.

Stanley, J. C. Rationale of the Study of Mathematically Precocious Youth (SMPY) during its first five years of promoting educational acceleration. In J. C. Stanley, W. C. George, & C. H. Solano (Eds.), The gifted and the creative: A fifty-year perspective. Baltimore, Md.: Johns Hopkins, 1977.